MASC: Metal-Aware Sampling and Correction via Reinforcement Learning for Accelerated MRI

Zhengyi Lu[1] ZHENGYI.LU@VANDERBILT.EDU
Ming Lu[2] MING.LU@VUMC.ORG
Chongyu Qu[1] CHONGYU.QU@VANDERBILT.EDU
Junchao Zhu[1] JUNCHAO.ZHU@VANDERBILT.EDU
Junlin Guo[1] JUNLIN.GUO@VANDERBILT.EDU
Marilyn Lionts[1] MARILYN.M.LIONTS@VANDERBILT.EDU
Yanfan Zhu[1] YANFAN.ZHU@VANDERBILT.EDU
Yuechen Yang[1] YUECHEN.YANG@VANDERBILT.EDU
Tianyuan Yao[1] TIANYUAN.YAO@VANDERBILT.EDU
Jayasai Rajagopal[3] RAJAGOPALJR@ORNL.GOV
Bennett Allan Landman[1] BENNETT.LANDMAN@VANDERBILT.EDU
Xiao Wang[3] WANGX2@ORNL.GOV
Xinqiang Yan[2] XINQIANG.YAN@VUMC.ORG
Yuankai Huo[1] YUANKAI.HUO@VANDERBILT.EDU

[1] *Vanderbilt University, Nashville, TN, USA 37215*
[2] *Vanderbilt University Medical Center, Nashville, TN, USA 37232*
[3] *Oak Ridge National Laboratory, Oak Ridge, TN, USA 37831*

Editors: Under Review for MIDL 2026

Abstract

Metal implants in MRI cause severe artifacts that degrade image quality and hinder clinical diagnosis. Traditional approaches address metal artifact reduction (MAR) and accelerated MRI acquisition as separate problems. We propose MASC, a unified reinforcement learning framework that jointly optimizes metal-aware k-space sampling and artifact correction for accelerated MRI. To enable supervised training, we construct a paired MRI dataset using physics-based simulation, generating k-space data and reconstructions for phantoms with and without metal implants. This paired dataset provides simulated 3D MRI scans with and without metal implants, where each metal-corrupted sample has an exactly matched clean reference, enabling direct supervision for both artifact reduction and acquisition policy learning. We formulate active MRI acquisition as a sequential decision-making problem, where an artifact-aware Proximal Policy Optimization (PPO) agent learns to select k-space phase-encoding lines under a limited acquisition budget. The agent operates on undersampled reconstructions processed through a U-Net-based MAR network, learning patterns that maximize reconstruction quality. We further propose an end-to-end training scheme where the acquisition policy learns to select k-space lines that best support artifact removal while the MAR network simultaneously adapts to the resulting undersampling patterns. Experiments demonstrate that MASC's learned policies outperform conventional sampling strategies, and end-to-end training improves performance compared to using a frozen pre-trained MAR network, validating the benefit of joint optimization. The code and models of MASC have been made publicly available: https://github.com/hrlblab/masc

Keywords: MRI reconstruction, k-space sampling, metal artifact reduction, reinforcement learning.

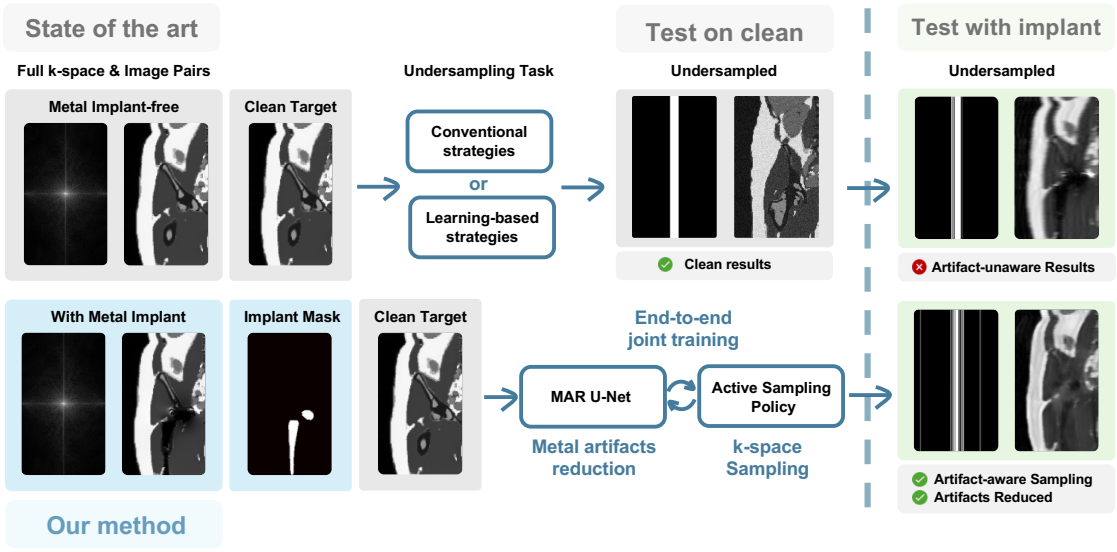

Figure 1: **Overview of MASC.** *Top:* Existing approaches using conventional or learning-based sampling strategies are trained on metal-free data and produce clean results when tested on implant-free cases, but fail when applied to metal-corrupted data, yielding artifact-unaware reconstructions. *Bottom:* MASC jointly optimizes a MAR U-Net and active sampling policy through end-to-end training on metal-corrupted data, enabling artifact-aware k-space sampling and effective artifact reduction.

1. Introduction

Magnetic resonance imaging (MRI) is a powerful diagnostic tool that provides excellent soft tissue contrast without ionizing radiation. However, two major challenges limit its clinical utility in certain patient populations: the presence of metal implants, which cause severe image artifacts due to susceptibility differences and signal voids (Hargreaves et al., 2011), and the inherently long scan times required to acquire fully-sampled k-space data (Liang and Lauterbur, 2000; Moratal et al., 2008; Zbontar et al., 2020). While both challenges have been extensively studied, existing approaches typically address them in isolation—metal artifact reduction (MAR) methods assume fully-sampled acquisitions, while accelerated MRI techniques focus on artifact-free subjects.

Metal implants are increasingly prevalent in patient populations, with hip replacements, spinal fusion hardware, and dental implants affecting millions of individuals annually (Maradit Kremers et al., 2015; Hegde et al., 2023). These metallic structures induce local magnetic field inhomogeneities that distort the MRI signal, manifesting as signal voids, geometric distortions, and bright streaking artifacts in reconstructed images (Koch et al., 2010; Hargreaves et al., 2011). Such artifacts can obscure adjacent anatomical structures critical for diagnosis, particularly in post-operative imaging where assessment of tissue surrounding the implant is essential. Accelerated MRI acquisition has emerged as a complementary research

direction, aiming to reduce scan times by acquiring only a subset of k-space lines (Wang et al., 2016; Han et al., 2019). Recent work has explored jointly optimizing fixed undersampling patterns with reconstruction networks (Zibetti et al., 2022), while reinforcement learning approaches frame k-space acquisition as a sequential decision-making problem, where an agent learns which phase-encoding lines to acquire to maximize reconstruction quality under a limited sampling budget (Zhang et al., 2019; Pineda et al., 2020; Bakker et al., 2020; Yen et al., 2024).

However, existing learned acquisition methods do not account for metal artifacts, and their interaction with artifact reduction remains unexplored. As illustrated in Figure 1, applying conventional active sampling policies to metal-corrupted data produces reconstructions that retain significant artifacts, since these policies were not designed to handle the unique k-space signatures induced by metal implants.

In this work, we propose MASC (Metal-Aware Sampling and Correction via Reinforcement Learning for Accelerated MRI), a unified framework that jointly addresses metal artifact reduction and accelerated MRI acquisition. Our approach formulates active k-space acquisition as a Markov decision process (Sutton and Barto, 2018), where an artifact-aware Proximal Policy Optimization (PPO) agent (Schulman et al., 2017) learns to select phase-encoding lines while a U-Net-based (Ronneberger et al., 2015) MAR network processes the undersampled reconstructions. To enable supervised training, we construct a paired dataset by generating digital phantoms from the AutoPET CT dataset (Gatidis et al., 2022) using multi-tissue segmentation, simulating MRI k-space data with and without virtual metal implants to produce exactly matched clean-metal pairs that provide ground truth supervision. Crucially, we introduce an end-to-end training scheme where both components co-adapt: the acquisition policy learns which k-space lines best support artifact removal, while the MAR network adapts to handle the specific undersampling patterns produced by the learned policy (Figure 1, bottom). This co-adaptation addresses a fundamental limitation of frozen MAR approaches—networks trained on fully-sampled data may not optimally handle the aliasing artifacts introduced by aggressive undersampling. We evaluate our approach against conventional sampling strategies and demonstrate that learned policies achieve superior reconstruction quality, with end-to-end training providing measurable improvements over frozen MAR baselines. Our contributions are:

- To the best of our knowledge, the first study to address accelerated MRI reconstruction in the presence of metal artifacts.
- A physics-based paired MRI dataset with simulated metal artifacts enabling supervised training for artifact reduction and reward computation for policy learning.
- An end-to-end artifact-aware reinforcement learning framework for joint optimization of k-space acquisition and metal artifact reduction.

2. Methods

2.1. Problem Formulation

We formulate active MRI acquisition as a Markov Decision Process (MDP) (Sutton and Barto, 2018). As shown in Figure 2, the state at time step t is defined as $s_t = (I_t, M_t)$,

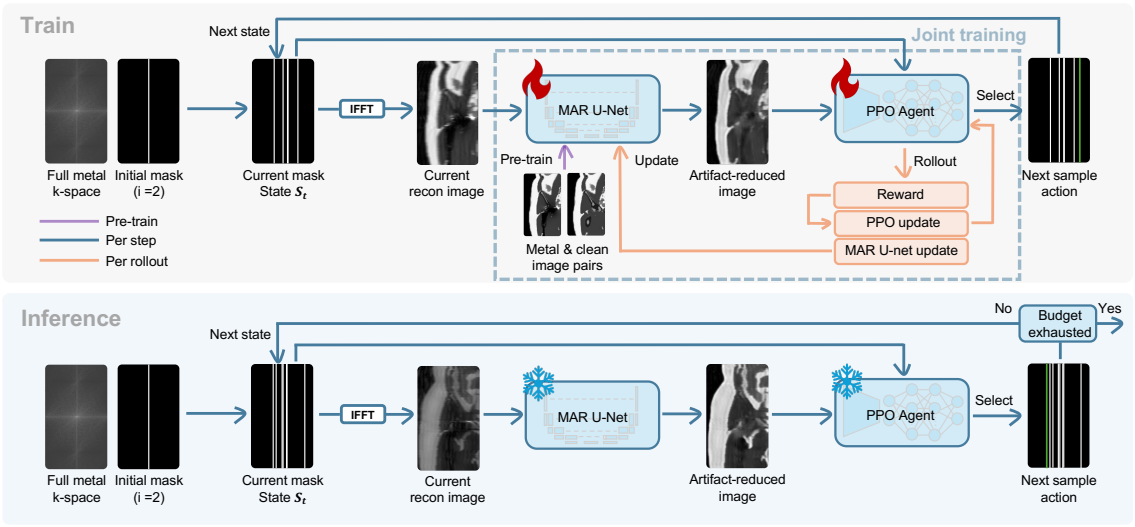

Figure 2: **MASC training and inference pipeline.** *Top (Train):* Starting from full metal-corrupted k-space, the current mask selects acquired lines for IFFT reconstruction. The MAR U-Net processes the reconstruction to produce artifact-reduced images, which the artifact-aware PPO agent observes to determine the next sampling action. Color-coded arrows indicate operation frequency: purple for pre-training (MAR U-Net on paired metal-clean images), blue for per-step operations (sequential k-space acquisition), and orange for per-rollout updates (reward computation, PPO policy update, and MAR U-Net fine-tuning). Fire icons denote trainable networks. *Bottom (Inference):* Both networks are frozen (snowflake icons). The agent iteratively selects k-space lines until the acquisition budget is exhausted.

where $I_t \in \mathbb{R}^{H\times W}$ represents the current magnitude reconstruction obtained from partially sampled k-space, and $M_t \in \{0,1\}^{N_{\text{pe}}}$ is a binary mask indicating acquired phase-encoding lines. The reconstruction is computed via inverse Fourier transform as $I_t = |\mathcal{F}^{-1}(K \odot M_t)|$, where $K \in \mathbb{C}^{H\times W}$ denotes the full k-space data (Liang and Lauterbur, 2000; Moratal et al., 2008).

The action space consists of selecting a single unacquired phase-encoding line. Upon selecting action a_t, the mask is updated as $M_{t+1}[a_t] = 1$, and the agent receives a reward based on reconstruction quality improvement:

$$r_t = \alpha \cdot (Q(I_{t+1}, I^*) - Q(I_t, I^*)) \tag{1}$$

where I^* is the ground truth image, α is a scaling factor, and Q measures reconstruction quality by combining Structural Similarity Index (SSIM) (Wang et al., 2004) and Normalized Mean Squared Error (NMSE):

$$Q(I, I^*) = \lambda_{\text{ssim}} \cdot \text{SSIM}(I, I^*) + \lambda_{\text{nmse}} \cdot (1 - \text{NMSE}(I, I^*)) \tag{2}$$

2.2. MASC Framework

Our framework comprises two components: an artifact-aware PPO-based (Schulman et al., 2017) artifact-aware acquisition policy and a U-Net-based MAR network (Ronneberger et al., 2015).

Acquisition Policy. The artifact-aware PPO agent uses a shared convolutional encoder for actor and critic branches. The encoder processes the concatenated reconstruction and acquisition mask, and the actor outputs action probabilities over unacquired lines with invalid actions masked. We optimize using the standard PPO clipped objective with GAE (Schulman et al., 2016).

MAR Network. The U-Net employs residual learning (He et al., 2016) where $g_\psi(I) = I + r_\psi(I)$, allowing the network to focus on learning artifact corrections rather than full image reconstruction.

Co-Adaptive Training. We propose a two-stage training scheme. In Stage 1, the MAR network is pre-trained on paired metal-corrupted and clean images using L1 and SSIM losses. In Stage 2, both networks are jointly optimized: the artifact-aware PPO agent observes MAR-processed reconstructions with rewards computed as:

$$r_t = \alpha \cdot (Q(g_\psi(I_{t+1}), I^*) - Q(g_\psi(I_t), I^*)) \tag{3}$$

where Q is our quality metric and I^* is the ground truth. This enables the policy to learn acquisition patterns complementary to the MAR network. Simultaneously, the MAR network is fine-tuned using:

$$\mathcal{L}_{\text{MAR}} = \|g_\psi(I_t) - I^*\|_2^2 \tag{4}$$

This co-adaptive training creates mutual influence: as the MAR network improves on particular acquisition patterns, the policy learns to favor them; as the policy converges, the MAR network specializes in those specific undersampling patterns. We alternate updates after each rollout with a lower learning rate for MAR fine-tuning. A detailed pseudo algorithm is provided in Appendix A.

3. Data and Experiments

3.1. Data

As shown in Figure 3, we construct a paired MRI dataset from 200 subjects in the AutoPET CT/PET dataset (Gatidis et al., 2022) using physics-based metal artifact simulation. This enables supervised training for artifact reduction and reward computation for policy learning. Both phantom segmentations and simulation results were validated by professional radiologists. Detailed subject information is provided in Appendix B.

Phantom Generation. For each CT volume, we run TotalSegmentator (Wasserthal et al., 2023) with three complementary tasks: (1) general segmentation for detailed organ delineation, (2) tissue composition for fat and muscle differentiation, and (3) body region for anatomical localization. These segmentations are combined to create multi-tissue 3D phantoms with tissue-specific MRI property maps including proton density, T1, and T2 relaxation times, following established conventions in the field (Segars et al., 2010; Bottomley et al., 1987; Pohmann et al., 2016; Rooney et al., 2007; Bojorquez et al., 2017; Gold et al.,

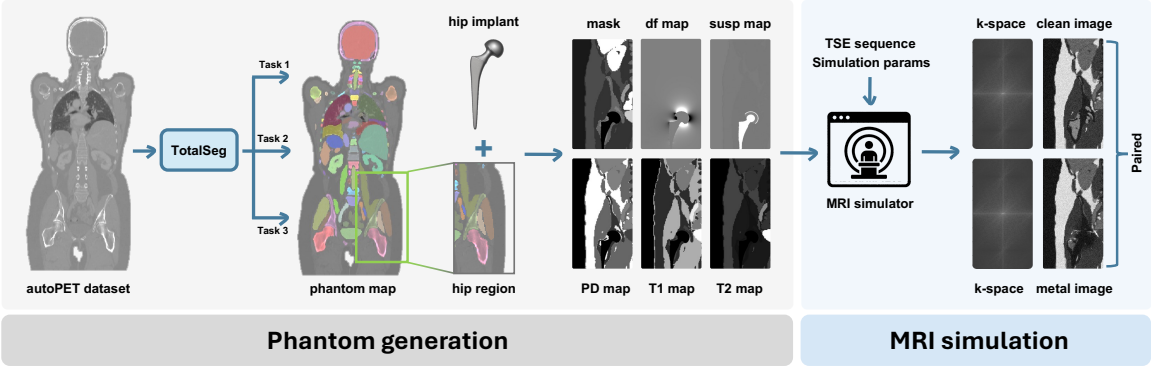

Figure 3: **Dataset construction pipeline.** *Phantom generation:* CT volumes from the autoPET dataset are processed through TotalSegmentator with three complementary tasks to produce multi-tissue phantom maps. The hip region is manually selected and combined with a cobalt-chromium hip implant model to generate tissue property maps including implant mask, off-resonance frequency (df), susceptibility (susp), proton density (PD), T1, and T2. *MRI simulation:* A physics-based simulator with TSE sequence parameters generates paired k-space data and reconstructions: clean images without metal and artifact-corrupted images with the virtual implant, providing exactly matched pairs for supervised training.

2004; Stanisz et al., 2005). We then manually select the hip region for each subject as the anatomically correct location for hip implant placement.

MRI Simulation. We employ the open-source MRI simulator developed by Zochowski et al. (Zochowski et al., 2024) to generate k-space data at 3T field strength. For metal artifact simulation, we use the total hip arthroplasty model provided in the simulator (Shi et al., 2017; Edelsbrunner et al., 1983; Mödinger et al., 2023), which comprises a femoral stem, femoral head, acetabular liner, and acetabular cup generated as spherical shells. The implant is modeled with cobalt-chromium (CoCr) alloy with magnetic susceptibility of 900 ppm, while tissue susceptibility values follow the simulator defaults: -9.05 ppm for soft tissues and water, -8.86 ppm for cortical bone, and -5.55 ppm for fat (Schenck, 1996; Smith et al., 2015; Mödinger et al., 2023). We manually fit the implant model into each subject's hip region with position adjustments to ensure anatomically realistic placement. The simulator models susceptibility-induced field inhomogeneities, signal dephasing, and geometric distortions. We simulate a Turbo Spin Echo (TSE) sequence with the following parameters: TR $= 4050$ ms, TE $= 32$ ms, readout bandwidth $= 710$ Hz/pixel, RF bandwidth $= 1$ kHz and slice thickness $= 3$ mm (Zochowski et al., 2024).

For each phantom, we generate: (1) clean k-space without metal, (2) metal-corrupted k-space with the virtual implant, and (3) a binary implant mask. Inverse Fourier transform yields paired reconstructions where each metal-corrupted image has an exactly matched clean reference serving as the ground truth target.

Final Dataset. The resulting dataset contains 200 subjects with 36 slices each, where edge slices are excluded to retain only the region affected by metal artifacts. This paired structure is critical for our framework: the clean reconstructions provide supervision for MAR network training, while the matched pairs enable reward computation during policy learning.

3.2. Experimental Setup

We partition the 200 subjects into training (160 subjects), validation (20 subjects), and test (20 subjects) sets, ensuring no subject overlap between splits. Each subject contains 36 slices, yielding 5,760 training, 720 validation, and 720 test slices. All experiments are conducted on an NVIDIA RTX 5090 GPU with 32 GB memory.

Training Configuration. The MAR network is pretrained for 100 epochs on paired metal-corrupted and clean images using the loss defined in Eq. (6). For end-to-end training, we use Adam optimizer (Kingma and Ba, 2015) with learning rate 3×10^{-4} for the artifact-aware PPO agent and 1×10^{-5} for the MAR network. PPO hyperparameters include: rollout length of 512 steps, 4 optimization epochs per update, clip range of 0.2, and entropy coefficient of 0.01. The reward scaling factor is $\alpha = 100$, and the quality metric weights are $\lambda_{\text{ssim}} = 0.7$ and $\lambda_{\text{nmse}} = 0.3$. All experiments use an acceleration factor of 10× (2 initial plus 18 budget acquired lines).

Baseline Methods. We compare against four conventional acquisition strategies and one learned baseline: (1) Center-out, acquiring lines from k-space center outward; (2) Random, uniform random selection; (3) Random-LowBias, random selection biased toward central k-space; (4) Equispaced, fixed equidistant sampling; and (5) SS-DDQN, a learned policy based on Double DQN for active MRI acquisition (Pineda et al., 2020). Conventional baselines are evaluated without MAR post-processing to demonstrate the advantage of our joint acquisition-and-reconstruction framework.

Ablation Study. We conduct ablation experiments with five configurations: (1) Clean-trained PPO without MAR; (2) Metal-trained PPO without MAR; (3) Clean PPO + Pretrained MAR (frozen); (4) Metal PPO + Pretrained MAR (frozen); and (5) MASC with joint co-adaptive optimization. The comparison between (4) and (5) isolates the benefit of end-to-end training versus applying MAR only at inference.

3.3. Evaluation Metrics

We assess reconstruction quality using SSIM, Peak Signal-to-Noise Ratio (PSNR), Mean Squared Error (MSE), NMSE, and Mean Absolute Error (MAE). All metrics are computed on final reconstructions after the acquisition budget is exhausted, reported as mean ± standard deviation across the test set.

4. Results

4.1. Comparison with Baseline Acquisition Strategies

Table 1 presents the quantitative comparison between our proposed MASC method and baseline acquisition strategies on the test set. Baseline methods are evaluated without MAR post-processing to isolate the effect of the acquisition strategy itself.

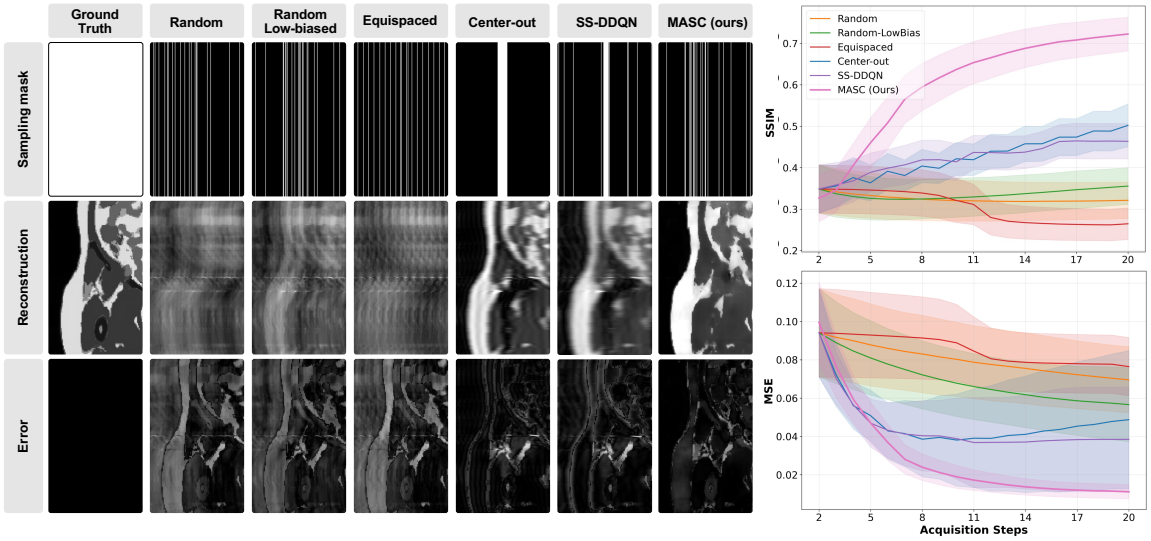

Figure 4: **Comparison of acquisition strategies.** *Left:* Ground truth, sampling masks, reconstructions, and error maps for each method including four conventional baselines (Random, Random Low-biased, Equispaced, Center-out), learned baseline (SS-DDQN), and our MASC. MASC produces reconstructions closest to ground truth with substantially darker error maps. *Right:* SSIM and MSE versus number of acquired k-space lines. Shaded regions indicate the std. MASC demonstrates superior performance throughout acquisition, with the gap widening as more lines are acquired.

Among conventional baselines, Center-out achieves the best performance, consistent with the understanding that low-frequency k-space components contain the majority of image energy and structural information. Random-LowBias outperforms uniform Random sampling by biasing selection toward the k-space center, while Equispaced sampling performs worst as it is designed for parallel imaging with coil sensitivity information unavailable in our single-coil setting. SS-DDQN, a learned acquisition policy based on Double DQN, outperforms Center-out in MSE and NMSE but underperforms in SSIM, suggesting that value-based RL can learn useful sampling patterns but may not fully capture perceptual quality objectives.

Our MASC method substantially outperforms all baselines across every metric, achieving 43.9% improvement in SSIM and 71.2% reduction in MSE compared to Center-out. This improvement can be attributed to two factors: the integration of MAR enables the policy to optimize for artifact-corrected reconstruction quality rather than raw undersampled images, and the end-to-end training allows the policy to discover acquisition patterns specifically tailored to the MAR network's correction capabilities.

Figure 4 provides qualitative comparison and acquisition dynamics. The left panel shows that Random and Random-LowBias produce severe aliasing due to incoherent sam-

Table 1: **Comparison with baseline acquisition strategies at 10× acceleration (2 initial + 18 budget lines).** Baselines are evaluated without metal artifact reduction. Results are mean ± std on the test set. Best results in **bold**. All improvements are statistically significant ($p < 0.001$, paired t-test).

Method	SSIM ↑	PSNR ↑	MSE ↓	NMSE ↓	MAE ↓
Random	0.3214 ± 0.0440	11.72 ± 1.11	0.0695 ± 0.0173	0.3374 ± 0.0929	0.2199 ± 0.0333
Random-LowBias	0.3556 ± 0.0444	12.70 ± 1.38	0.0566 ± 0.0193	0.2696 ± 0.0718	0.1877 ± 0.0332
Equispaced	0.2648 ± 0.0380	11.26 ± 0.92	0.0764 ± 0.0152	0.3716 ± 0.0846	0.2331 ± 0.0297
Center-out	0.5022 ± 0.0519	14.23 ± 3.17	0.0488 ± 0.0360	0.2249 ± 0.1352	0.1507 ± 0.0619
SS-DDQN	0.4637 ± 0.0425	14.90 ± 2.41	0.0385 ± 0.0273	0.1798 ± 0.1017	0.1355 ± 0.0460
MASC (Ours)	**0.7224 ± 0.0408**	**19.73 ± 1.28**	**0.0111 ± 0.0036**	**0.0546 ± 0.0190**	**0.0640 ± 0.0121**

pling, while Equispaced results in structured aliasing patterns. Center-out preserves gross anatomy but loses fine details from missing high-frequency information. SS-DDQN shows improved structure but retains visible artifacts. In contrast, MASC produces reconstructions closest to ground truth with substantially darker error maps.

The right panel reveals the temporal dynamics of acquisition. MASC exhibits a steeper improvement trajectory in early steps, suggesting the learned policy prioritizes k-space lines with maximal information gain—a behavior consistent with optimal experimental design principles where early measurements should maximize expected information. The widening performance gap as acquisition progresses indicates that MASC's sequential decisions compound beneficially, whereas baseline methods lack this adaptive capability. Notably, MASC maintains consistently lower variance (narrower shaded regions), demonstrating robust performance across diverse anatomical structures and artifact patterns.

4.2. Ablation Study

We conduct an ablation study to analyze the contribution of each component in our proposed MASC framework. Table 2 presents results of five configurations that systematically evaluate the effects of training data, MAR integration, and end-to-end optimization. Further visualizations are shown in Appendix C.

Comparing Clean-trained PPO and Metal-trained PPO reveals how metal artifacts affect policy learning without MAR post-processing. Clean-trained PPO achieves higher SSIM while Metal-trained PPO yields lower MSE, suggesting a trade-off: artifact-free training preserves structural learning, but metal-aware training enables partial adaptation to the artifact distribution. This can be understood through the lens of reward signal quality—metal artifacts corrupt the reconstruction used for reward computation, effectively adding noise to the policy gradient and hindering convergence to optimal sampling strategies.

Adding pretrained MAR substantially improves all configurations, with Metal PPO + Pretrained MAR slightly outperforming its clean-trained counterpart. This suggests that when MAR corrects artifacts at inference, the metal-trained policy's implicit knowledge of artifact locations becomes beneficial rather than detrimental. The policy may learn to sam-

Table 2: **Ablation study on training strategy components at 10× acceleration.** Configurations progressively add metal-aware training, MAR integration, and end-to-end optimization. Results are mean ± std on the test set. Best results in **bold**. MASC significantly outperforms all ablation variants ($p < 0.001$, paired t-test).

Configuration	SSIM ↑	PSNR ↑	MSE ↓	NMSE ↓	MAE ↓
Clean-trained PPO	0.5120 ± 0.0559	14.47 ± 3.19	0.0463 ± 0.0348	0.2137 ± 0.1309	0.1465 ± 0.0606
Metal-trained PPO	0.4758 ± 0.0465	14.99 ± 2.47	0.0373 ± 0.0220	0.1766 ± 0.0926	0.1341 ± 0.0425
Clean PPO + Pretrained MAR	0.6640 ± 0.0600	17.08 ± 1.68	0.0212 ± 0.0095	0.1020 ± 0.0400	0.0945 ± 0.0262
Metal PPO + Pretrained MAR	0.6874 ± 0.0621	17.72 ± 1.59	0.0182 ± 0.0077	0.0888 ± 0.0384	0.0820 ± 0.0213
MASC (Ours)	**0.7224 ± 0.0408**	**19.73 ± 1.28**	**0.0111 ± 0.0036**	**0.0546 ± 0.0190**	**0.0640 ± 0.0121**

ple k-space regions that, while suboptimal for direct reconstruction, provide complementary information for the MAR network's correction process.

Our proposed MASC framework achieves the best performance across all metrics, improving SSIM by 5.1% and reducing MSE by 39.0% compared to the strongest baseline. This improvement stems from breaking the train-test distribution mismatch: in post-hoc approaches, the MAR network is trained on fully-sampled data but applied to undersampled reconstructions with aliasing artifacts it has never encountered. End-to-end training resolves this by exposing the MAR network to the actual undersampling patterns produced by the policy, while simultaneously allowing the policy to discover acquisition strategies that maximize the MAR network's correction capability. The consistently lower standard deviation across all metrics further indicates that this co-adaptation produces more robust reconstructions, as both components converge to a mutually beneficial equilibrium rather than operating independently.

5. Conclusion

We presented MASC, a unified reinforcement learning framework that jointly optimizes metal-aware k-space sampling and artifact correction for accelerated MRI. Our approach contributes a physics-based paired MRI dataset with simulated metal artifacts from 200 subjects, enabling supervised training for artifact reduction and reward computation for policy learning, along with an end-to-end co-adaptive training scheme where the PPO-based acquisition policy and U-Net-based MAR network mutually adapt to each other. Experiments demonstrate that MASC substantially outperforms both conventional and learned acquisition strategies, achieving 43.9% improvement in SSIM and 77.3% reduction in MSE compared to the best conventional baseline at 10× acceleration. Ablation studies confirm that end-to-end training provides measurable benefits over using a frozen pre-trained MAR network, validating the importance of joint optimization.

The current limitation of our work is the focus on a single implant type (cobalt-chromium hip implant) at one anatomical region. Future work will extend MASC to diverse implant materials with varying magnetic susceptibilities, different anatomical regions such as spine and knee, and multi-coil settings where parallel imaging can be combined with learned acquisition policies.

Acknowledgments

This research was supported by NIH R01DK135597 (Huo), DoD HT9425-23-1-0003 (HCY), NSF 2434229 (Huo) and KPMP Glue Grant. This work was also supported by Vanderbilt Seed Success Grant and Vanderbilt-Liverpool Seed Grant. This research was also supported by NIH grants R01EB033385, R01DK132338. We extend gratitude to NVIDIA for their support by means of the NVIDIA hardware grant. This research was also supported by NIH R01 EB031078 (Yan), R21 EB029639 (Yan), R03 EB034366 (Yan). This manuscript has been co-authored by ORNL, operated by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S.Department of Energy.

References

Tim Bakker, Herke van Hoof, and Max Welling. Experimental design for MRI by greedy policy search. In *Advances in Neural Information Processing Systems*, volume 33, pages 18954–18966, 2020.

Josefina Z Bojorquez, Stéphanie Bricq, Clement Acquitter, et al. What are normal relaxation times of tissues at 3T? *Magnetic Resonance Imaging*, 35:69–80, 2017.

Paul A Bottomley, Christopher J Hardy, Robert E Argersinger, et al. A review of 1H nuclear magnetic resonance relaxation in pathology: Are T1 and T2 diagnostic? *Medical Physics*, 14(1):1–37, 1987.

Herbert Edelsbrunner, David Kirkpatrick, and Raimund Seidel. On the shape of a set of points in the plane. *IEEE Transactions on Information Theory*, 29(4):551–559, 1983. doi: 10.1109/TIT.1983.1056714.

Sergios Gatidis et al. The AutoPET challenge: Towards fully automated lesion segmentation in oncologic PET/CT imaging. *Research Square Preprint*, 2022.

Garry E Gold, Eric Han, Jeff Stainsby, Graham Wright, Jean Brittain, and Christopher Beaulieu. Musculoskeletal MRI at 3.0 T: Relaxation times and image contrast. *American Journal of Roentgenology*, 183(2):343–351, 2004. doi: 10.2214/ajr.183.2.1830343.

Yoseob Han, Leonard Sunwoo, and Jong Chul Ye. k-space deep learning for accelerated MRI. *IEEE Transactions on Medical Imaging*, 39(2):377–386, 2019.

Brian A Hargreaves, Pauline W Worters, Kim Butts Pauly, John M Pauly, Kevin M Koch, and Garry E Gold. Metal-induced artifacts in MRI. *American Journal of Roentgenology*, 197(3):547–555, 2011. doi: 10.2214/AJR.11.7364.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, 2016.

Vivek Hegde, Jeffrey B Stambough, Brett R Levine, and Bryan D Springer. Highlights of the 2022 American Joint Replacement Registry annual report. *Arthroplasty Today*, 21: 101137, 2023. doi: 10.1016/j.artd.2023.101137.

Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2015.

Kevin M Koch, Brian A Hargreaves, Kim Butts Pauly, Weitian Chen, Garry E Gold, and Kevin F King. Magnetic resonance imaging near metal implants. *Journal of Magnetic Resonance Imaging*, 32(4):773–787, 2010. doi: 10.1002/jmri.22313.

Zhi-Pei Liang and Paul C Lauterbur. *Principles of Magnetic Resonance Imaging: A Signal Processing Perspective*. SPIE Optical Engineering Press, 2000.

Hilal Maradit Kremers, Dirk R Larson, Cynthia S Crowson, et al. Prevalence of total hip and knee replacement in the united states. *Journal of Bone and Joint Surgery*, 97(17): 1386–1397, 2015. doi: 10.2106/JBJS.N.01141.

Yvonne Mödinger, Eric D Anttila, Graeme M Baker, David C Gross, and Alessandro A Porporati. Magnetic resonance safety evaluation of a novel alumina matrix composite ceramic knee and image artifact comparison to a metal knee implant of analogous design. *Arthroplasty Today*, 22:101170, 2023. doi: 10.1016/j.artd.2023.101170.

David Moratal, Amparo Vallés-Luch, Luis Martí-Bonmatí, and Marijn E Brummer. k-space tutorial: An MRI educational tool for a better understanding of k-space. *Biomedical Imaging and Intervention Journal*, 4(1), 2008.

Luis Pineda, Sumana Basu, Adriana Romber, Francesco Caliva, and Christian Bluethgen. Active mr k-space sampling with reinforcement learning. In *MICCAI*, pages 23–33, 2020.

Rolf Pohmann, Oliver Speck, and Klaus Scheffler. Signal-to-noise ratio and MR tissue parameters in human brain imaging at 3, 7, and 9.4 tesla using current receive coil arrays. *Magnetic Resonance in Medicine*, 75(2):801–809, 2016.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *MICCAI*, pages 234–241, 2015.

William D Rooney, Glyn Johnson, Xin Li, et al. Magnetic field and tissue dependencies of human brain longitudinal 1H2O relaxation in vivo. *Magnetic Resonance in Medicine*, 57 (2):308–318, 2007.

John F Schenck. The role of magnetic susceptibility in magnetic resonance imaging: MRI magnetic compatibility of the first and second kinds. *Medical Physics*, 23(6):815–850, 1996.

John Schulman, Philipp Moritz, Sergey Levine, Michael Jordan, and Pieter Abbeel. High-dimensional continuous control using generalized advantage estimation. *arXiv preprint arXiv:1506.02438*, 2016.

John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.

W Paul Segars, Gregory Sturgeon, S Mendonca, et al. 4d xcat phantom for multimodality imaging research. *Medical Physics*, 37(9):4902–4915, 2010.

Xinwei Shi, Daehyun Yoon, Kevin M Koch, and Brian A Hargreaves. Metallic implant geometry and susceptibility estimation using multispectral B0 field maps. *Magnetic Resonance in Medicine*, 77(6):2402–2413, 2017.

Matthew R Smith, Nathan S Artz, Curtis Wiens, Diego Hernando, and Scott B Reeder. Characterizing the limits of MRI near metallic prostheses. *Magnetic Resonance in Medicine*, 74(6):1564–1573, 2015. doi: 10.1002/mrm.25540.

Greg J Stanisz, Ewa E Odrobina, Joseph Pun, et al. T1, T2 relaxation and magnetization transfer in tissue at 3T. *Magnetic Resonance in Medicine*, 54(3):507–512, 2005. doi: 10.1002/mrm.20605.

Richard S Sutton and Andrew G Barto. *Reinforcement Learning: An Introduction*. MIT Press, 2018.

Shanshan Wang et al. Accelerating magnetic resonance imaging via deep learning. In *Proc. IEEE 13th Int. Symp. Biomed. Imag. (ISBI)*, pages 514–517, 2016.

Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004.

Jakob Wasserthal, Hanns-Christian Breit, Manfred T Meyer, Maurice Pradella, Daniel Hinck, Alexander W Saber, et al. TotalSegmentator: Robust segmentation of 104 anatomic structures in CT images. *Radiology: Artificial Intelligence*, 5(5):e230024, 2023.

Chen-Yu Yen, Raghav Singhal, Umang Sharma, Rajesh Ranganath, Sumit Chopra, and Lerrel Pinto. Adaptive sampling of k-space in magnetic resonance for rapid pathology prediction, 2024.

Jure Zbontar, Florian Knoll, Anuroop Sriram, Matthew J Muckley, Mary Bruno, Aaron Defazio, Marc Parente, Krzysztof J Geras, Joe Katsnelson, Hersh Chandarana, Zizhao Zhang, Michal Drozdzal, Adriana Romero, Michael Rabbat, Pascal Vincent, James Pinkerton, Duo Wang, Nafissa Yakubova, Erich Owens, C Lawrence Zitnick, Michael P Recht, Daniel K Sodickson, and Yvonne W Lui. fastMRI: An open dataset and benchmarks for accelerated MRI. *Radiology: Artificial Intelligence*, 2(6):e190007, 2020.

Zizhao Zhang, Adriana Romero, Matthew J Muckley, Pascal Vincent, Lin Yang, and Michal Drozdzal. Reducing uncertainty in undersampled MRI reconstruction with active acquisition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2049–2058, 2019.

Marcelo Victor Wust Zibetti, Florian Knoll, and Ravinder R Regatte. Alternating learning approach for variational networks and undersampling pattern in parallel MRI applications. *IEEE Transactions on Computational Imaging*, 8:449–461, 2022.

Kevin C Zochowski et al. Open-source simulator of imaging near metal at arbitrary magnetic field strengths. *Magnetic Resonance in Medicine*, 2024.

Appendix A. Additional Implementation Details

Algorithm A outlines our co-adaptive training procedure, where the PPO policy and MAR-UNet are jointly optimized after each rollout.

Algorithm 1: MASC: Joint PPO and MAR-UNet Training

Input : Dataset $\mathcal{D}$, rollout length N, scaling factor α
Output: Trained policy π_θ and MAR-UNet g_ψ
Initialize policy π_θ, value network V_ϕ, pretrained MAR-UNet g_ψ
while *not converged* **do**
  Initialize rollout buffer $\mathcal{B} \leftarrow \emptyset$
  foreach *step n in rollout* **do**
    Reset environment if episode done
    Sample action $a \sim \pi_\theta(\cdot|s)$
    Update mask $m' \leftarrow m \cup \{a\}$
    Reconstruct $I_{\text{recon}} \leftarrow \mathcal{F}^{-1}(k_{\text{full}} \odot m')$
    Apply MAR-UNet $\hat{I} \leftarrow g_\psi(I_{\text{recon}})$
    Compute reward $r \leftarrow \alpha \cdot (Q(\hat{I}, I^*) - Q(\hat{I}_{\text{prev}}, I^*))$
    Store transition in $\mathcal{B}$
  end
  Update PPO policy θ and value ϕ using $\mathcal{B}$
  Update MAR-UNet ψ with $\mathcal{L}_{\text{MAR}} = \|g_\psi(I_t) - I^*\|_2^2$
end
return π_θ, g_ψ

Appendix B. Additional Dataset Details

We selected 200 subjects from the AutoPET FDG-PET/CT dataset (Gatidis et al., 2022) for MRI simulation. Table 3 summarizes patient demographics.

Table 3: **Dataset demographics (n=200).** Summary of patient characteristics from the AutoPET dataset used for MRI simulation.

Characteristic	**Value**
Subjects	200
Age (years)	62.0 ± 15.4 (range: 11–95)
Sex	
Male	115 (57.5%)
Female	84 (42.0%)
Not reported	1 (0.5%)
Diagnosis	
Melanoma	70 (35.0%)
Lung cancer	65 (32.5%)
Lymphoma	61 (30.5%)
Negative	4 (2.0%)

Appendix C. Additional Ablation Study Visualization

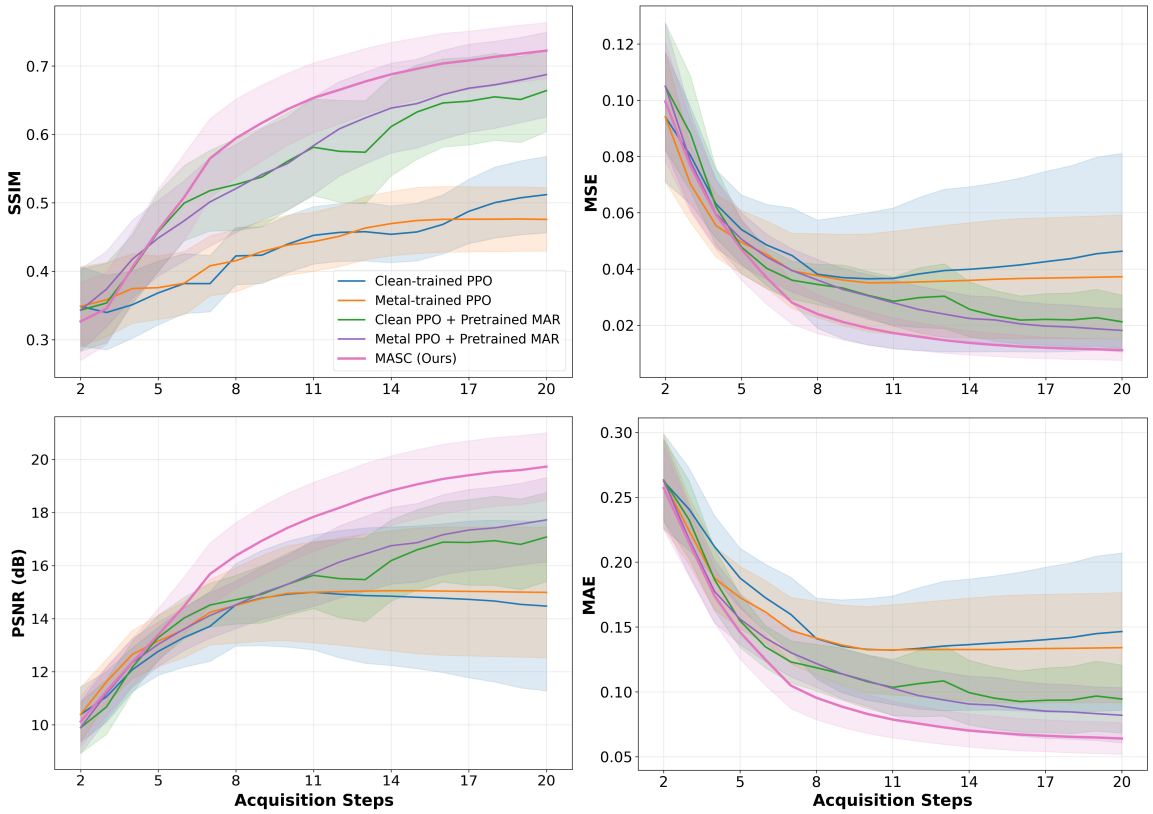

Figure 5: **Ablation study on training strategy components at 10× acceleration.** Configurations progressively add metal-aware training, MAR integration, and end-to-end optimization. SSIM, MSE, PSNR and MAE versus number of acquired k-space lines. Shaded regions indicate the std. MASC demonstrates superior performance throughout acquisition, with the gap widening as more lines are acquired.